

ثم توفير هذه النسخة من طرف الطالب علي أعشرين

"الجغرافي هو المواطن الحقيقي في جميع التخصصات ، قد يكون الشاعر شاعرا رومنسيا ، وقد يكون الفرنسي مبدعا ، لكن الجغرافي بالتميز هو سيد الجغرافيا والجغرافيا هي المجال والوطن ، ودائما ما يبكي في محاربه الخاصة ."

كلية الآداب والعلوم الإنسانية عين الشق

قواعد الاحصاء

د. سعيد اجديرا
مقرر مجزوءة الإحصاء الفصل الثاني شعبة الجغرافيا
كلية الآداب والعلوم الإنسانية عين الشق
جامعة الحسن الثاني
الدار البيضاء

الإحصاء و التعبير البياني

المحاور الأساسية لمادة الإحصاء:

البرنامج العام للمادة ينقسم إلى قسمين متوازيين: نظري وتطبيقي.

القسم النظري:

تعاريف وخصائص وملاحظات وأمثلة تطبيقية، ضرورية كلها لاستيعاب المكونات الأساسية للجانب التطبيقي.

1) تقديم عام.
2) الإحصائيات ذات متغير واحد: مقاييس النزعة المركزية - مقاييس التشتت.
3) الإحصائيات ذات متغيرين.
4) تحليل العلاقات والارتباطات بين المتغيرات الإحصائية.
5) الارتباط والانحدار والتوفيق الخطي.
6) السلاسل الزمنية والتوقعات المستقبلية.
7) الأرقام الاستدلالية والتوقعات الاقتصادية.

القسم التطبيقي:

بحوث للطلبة تُنجز تدريجيا خلال السنة الدراسية، بالموازاة مع إنجاز القسم النظري، وكذلك مع تهيئ أعمال ميدانية. وهي دراسات ليست وصفية فحسب، بل تحليلية أيضا. لكنها بالخصوص تطبيقية في عدة مواضيع، تتماشى غالبا مع القضايا التي يهتم بها المسلك الجامعي. فتكون مواضيع العروض والبحوث ذات فائدة علمية متجددة، لأن معلوماتها وبياناتها الإحصائية تتجدد باستمرار، مثلما هي في سائر العلوم الإنسانية.

ومن المواضيع المقترحة، التي قد تصلح لبحوث جامعية:
1) الإحصاء: تعريفه العام والخاص حسب ميادين تطبيقه.
2) مصادر الإحصاء وتقنيات الاستمارات فيه.
3) الإحصاء الجغرافي: تعريف ونماذج تطبيقية.
4) الإحصاء التاريخي: تعريف ونماذج تطبيقية.
5) الإحصاء وتطبيقاته في العلوم الإنسانية.
6) الإحصاء الاجتماعي: تعريف ونماذج تطبيقية.
7) الإحصاء التربوي: تعريف ونماذج تطبيقية.
8) الإحصاء النفسي: تعريف ونماذج تطبيقية.
9) الإحصاء والتوقعات المستقبلية.
10) الإحصاء وقضية تعريب المصطلحات.
11) استعمال الآلة الحاسبة في الإحصاء.
12) استعمال المعلوميات والحواسيب في الإحصاء.
13) الإحصاء والجودة في العلوم الاقتصادية.

<u>المراجع:</u> لتنمية المعلومات الإحصائية.

متعددة ومتنوعة، إلا أن مما قد يتماشى منها مع مواضيع مذكورة ما يلي:

1) الإحصاء الوصفي مع حلول لتمارين تطبيقية ـ مبارك بلمغنية.
2) معجم مصطلحات الإحصاء والاحتمالات ـ مصطفى بنيخلف.
3) الإحصاء الجغرافي – فتحي عبد الله فياض.
4) التحليل الإحصائي للبيانات الجغرافية - فتحي عبد الله فياض.
5) الإحصاء والبحث التاريخي – مصطفى زايد.
6) علم الإحصاء وتطبيقاته في المجالين التربوي والاجتماعي ـ حمودي.
7) مقدمة طرق الإحصاء الاجتماعي ـ الهانسي.
8) الإحصاء الاجتماعي ـ بلاكوك.
9) الإحصاء الوصفي والاستدلالي في العلوم النفسية والتربوية ـ منسي.
10) الإحصاء التربوي ـ المركز العربي.
11) الإحصاء التربوي ـ عفانة.
12) الإحصاء في التربية ـ رشيد.
13) الإحصاء السيكولوجي التطبيقي ـ عيسوي.
14) الإحصاء النفسي وقياس القدرات الإنسانية ـ أسعد.
15) الإحصاء في علم النفس ـ فرج.
16) مناهج البحث وطرق التحليل الإحصائي – فؤاد أبو حطب.
17) Pratiques statistiques en sciences humaines et sociales – Longouet.
18) Méthodes statistiques en sciences humaines – Howell
19) Statistique en sciences humaines – Rouanet.
20) دليل الأطروحات والرسائل الجامعية.

الإحصاء الوصفي:

1) تقديم له:

أصله باعتباره تقنيات حسابية لدراسة المجموعات، قديم قدم وجود دول منظمة، لأن الإحصائيات ضرورية لتسيير كل دولة. فكان اسما على مسمى: state – état – status ...

و الميادين التي يُستعمل فيها متعددة ومتنوعة المجالات، من أهمها ما هو معاصر. فكان منها:

+ الجغرافي ذو البعد المكاني (الدول – صبيب المياه ...)

+ المناخي (الأمطار – الرياح ...)

+ الفلاحي (التربة – المزروعات ...)

+ الديموغرافي (السكان - المواليد – الوفيات ...)

+ الاجتماعي (الأسر – الهجرة – السكن – الفقر – البطالة ...)

+ الاقتصادي (التجارة/الصناعة – الإنتاج/الاستهلاك – الدخل ...)

و بفضل تطور الإعلاميات والحواسيب، تطور علم الإحصاء كثيرا، إذ انضبط لأقصى حد.

2) مراحله:

الأولى: جمع وضبط المعلومات اللازمة في دراسة إحصائية، عبر تصنيف وتبويب البيانات الإحصائية. وهذه المعطيات تكون مُبوبة، أو غير مبوبة بجدول. والتبويب يكون أحادي المدخل بظاهرة مدروسة واحدة، أو متعدد المداخل في حالة ظاهرتين فما فوق. فيكون الجدول الإحصائي، إما بسيطا أو مركبا.

الثانية: تحليل المعلومات عبر طرائق إحصائية محددة.

الثالثة: استخراج الاستنتاجات، خاصة تلك التي تستهدفها الدراسة الإحصائية.

3) الرمز عامل المجموع Σ : (sigma): مشتق من خصائص الجمع. فكان:

$$\sum_{i=1}^{n} xi = x_1 + x_2 + \ldots + x_n$$

تمرين:

x_i	1	3	6	7	9
n_i	5	3	1	4	2
y_i	0	1	2	3	4

احسب مباشرة أو بجدول:

$\sum_{i=1}^{n} xi$ ، $N = \sum_{i=1}^{n} ni$ ، $\sum_{i=1}^{n} yi$ ، $\sum_{i=1}^{n} xi * ni$ ، $\sum_{i=1}^{n} ni * xi^2$ ، $m = \sum_{i=1}^{n} xi * ni \,/\, N$ ،

$\sum_{i=1}^{n} ni / xi - m /$ ، $\sum_{i=1}^{n} ni(xi - m)^2$ ، $\sum_{i=1}^{n} ni(xi - m) = 0$ ، $\sum_{i=1}^{n} (xi - yi)^2$.

الإحصائيات ذات متغير واحد.

I) أمثلة:

هناك ثلاثة أنواع حسب طبيعة المتغير الإحصائي المدروس، لأنها تكون إما نوعية أو كمية، والكمية تكون إما متصلة أو متقطعة (متقطعة كنقط قسم مثلا ومتصلة كالسن).

مثال 1: كيفي.

الجدول التالي يعطي عدد الأجانب المقيمين بالمغرب حسب إحصاء 1960

الجنسية	عدد الأفراد التكرار	التكرار المتراكم	التردد	التردد المتراكم	التردد المئوي %
فرنسيون	175090	175090	0,44	0,44	44
إسبان	92901	267991	0,23	0,67	23
جزائريون	93026	361017	0,24	0,91	24
آخرون	34866	395883	0,09	1,00	9

مثال 2: كمية متقطعة.

الجدول التالي يعطي عدد الحوادث اليومية في مدينة معينة لمدة 50 يوما.

عدد الحوادث	عدد الأيام
0	21
1	18
2	7
3	3
4	1

مثال 3: كمية متصلة.

أجريت تجربة على 400 مصباح كهربائي لتحديد مدة صلاحيته.

مدة الصلاحية بمئات الساعات	عدد المصابيح
[3,4[	15
[4,5[	46
[5,6[	54
[6,7[	78
[7,8[	70
[8,9[	64
[9,10[	45
[10,11[	20
[11,12[	8

II) مصطلحات:

statistique descriptive :الإحصاء الوصفي
statistique analytique / mathématique :الإحصاء التحليلي
population / échantillon :الساكنة / عينة
individu statistique :وحدة إحصائية (= فرد إحصائي)
caractère (Variable) :الميزة (المتغير الإحصائي)
modalité :قيمة الميزة (= مواصفة)
série statistique :متسلسلة إحصائية
séries / distributions :سلاسل / توزيعات
série chronologique (= chronique) :متسلسلة تاريخية (= زمنية)
qualitative / quantitative :كيفية (= نوعية) / كمية
continue / discontinue = discrète :متصلة / متقطعة
simple / pondérée :بسيطة / متزنة (= مُرجَّحة)
intervalle = classe :مجال (= صنف = فئة = فصيلة مجالية)
centre = milieu :مركز (= وسط = منتصف)
effectif (global / partiel :الحصيص (الإجمالي / الجزئي)
fréquence (cumulée / partiel :تردد (متراكم / جزئي)
croissant / décroissant :تزايدي = تصاعدي / تناقصي = تنازلي
pourcentage :النسبة المئوية
tableaux statistiques :الجداول الإحصائية
représentations graphiques :التمثيلات المبيانية
diagramme :مخطط
en bâtons :بالعصي (= عصوي)
(à bandes (= barres :بالأشرطة (= بالقضبان)
)sectoriel (circulaire :قطاعي (دائري)
polygone :مضلع
histogramme :مدراج
pyramide :هرم
paramètres de position :وسيطات الوضع (مقاييس التموضع)
)mode (classe modale :منوال (صنف منوالي)
moyenne arithmétique :المعدل (الوسط) الحسابي
géométrique / harmonique :الهندسي / التوافقي
quadratique :المعدل الأربوعي (التربيعي)
valeur médiane / médiale :قيمة وسطية = الوسيط
paramètres de dispersion :وسيطات التشتت
range/étendue :المدى (الفارق المطلق)
écart moyen / type :الانحراف المتوسط (الفارق المطلق المتوسط)
écart type / variance :الانحراف الطرازي / المغايرة (التباين)
coefficient de variation :معامل التغير (الاختلاف)
coefficient d’asymétrie / de concentration :معامل الالتواء / التمركز

III) تعاريف وأمثلة:

1 - المجموعة الإحصائية Population:

هي المجموعة التي تخضع للدراسة، كيفما كانت مكونة من أفراد.

2 - فرد إحصائي Individu statistique:

كل عنصر من المجموعة الإحصائية، ككل فرد من سكان منطقة ما.

3 – أفراد أو مجموعة عينة Echantillon:

جزء من المجموعة الإحصائية يمثلها، خصوصا عندما تكون ساكنة كبيرة. فتتم معاينة وعينة إحصائية، يُطلب فيها أخذ المجموعة n بحوالي 10% من المجتمع الإحصائي N، إذ توضع نسبة الخطأ المسموح به عند تحديد مجموعة. ويلزم على العموم، أن تكون المجموعة مناسبة لتمثل مجتمعها كما يجب.

4 – المتغير الإحصائي أو الميزة الاحصائية Caractère Variable:

هي موضوع الدراسة الإحصائية، إلا أنها نوعان:

* **كمية:** يمكن التعبير عنها بأعداد، كعدد الأولاد ومجالات كالسن والطول بالسنتمتر مثلا ودرجة الحرارة المئوية المصنفة.

* **نوعية:** ليست كذلك، كالجنسية واللون والجهات...

فيرمز للكمية ب Xi، أو بأصناف مرتبة حسب طبيعة قيم المتغير الاحصائي. لكن الفئات يكون لها نفس الطول متتابعة، وإلا عُدّلت.

5 - التكرارni:

هو عدد تكرار المواصفة، أو الموافق للنوعية. فيكون بمعنى التكرار عددا طبيعيا صحيحا، دالا على التكرر المنفصل **في** المتغير الاحصائي، لتفادي أمثلة خاطئة إحصائيا سواء من حيث المعطيات أو الوسيطات والمقاييس، كإحصاء فرد عدة مرات في مواقع مختلفة للمتغير الاحصائي مثل طلبة يُحصون في عدة وحدات، أو حيث لا يُستساغ التكرار كالسنوات لأنها لا تتكرر في ذاتها.

و المجموعة {(xi,ni)} تكون متسلسلة إحصائية ذات متغير واحد، بسيطة أحيانا، لكنها مرتبة قطعا. ويمكن مراكمة الحصيصات تصاعديا في الغالب، أو تنازليا، مع الرمز N i.

6 – التكرار الإجمالي:

المرموز له بـ N، هو مجموع الحصيصات أي كل تكرارات قيم المتغير الاحصائي.

7 – التردد:

هو خارج قسمة التكرار على التكرار الإجمالي، ويرمز له بـ fi. فيكون $fi = ni / N$. لكن حتما $\sum fi = 1$، إذ يمكن مراكمة الترددات تزايديا في الغالب، أو تنازليا مع الرمز F i.

8 – التردد المئوي:

هي جداء التردد في مائة، أي $pi = fi*100\ \%$. لكن حتما $\sum pi = 100\ \%$.

9 – ملائمة الفئات:

إذا كان التوزيع غير منتظم، أي غير متساوي طول الفئات.

يُستخدم التكرار المعدل أو كثافة التكرار، بقسمة كل تكرار فئة على طولها. فتضاف خانة التكرارات المعدلة بإفراط، مع مراعاة التكرار الإجمالي.

و في حالة قلة عدد الفئات غير المنتظمة، يكفي استعمال القاعدة الثلاثية لتحديد التكرار المعدل: طول الفئة ← تكرارها

الطول المنتظم ← التكرار المعدل

مثال 1: في إحصائيات أساسية لسنة 2001 حول نِسب فئات عُمرية:

أقل من 15 سنة: 6 , 31 %. من 15 إلى 59 سنة: 61 %. أكثر من 60 سنة: 4 , 7 %.

ملاحظة: تُزال هنا الفاصلة بضرب الكل في 10، فيصير التكرار الإجمالي 1000. ثم يُعـدّل الجدول، لأن الفئات غير متقايسـة، مع اعتبار سـن أقصى هو 120.

مثال 2:

قيمة عدد الحوادث xi	التكرار عدد الأيام ni	التكرار المتراكم Ni	التردد fi	التردد المتراكم Fi	التردد المئوي pi
0	21	21	0,42	0,42	42
1	18	39	0,36	0,78	36
2	7	46	0,14	0,92	14
3	3	49	0,06	0,98	6
4	1	50	0,02	1	2

مثال 3:

مدة الصلاحية	مركز الصنف	التكرار	التكرار المتراكم	التردد	التردد المئوي
الصنف / المجال		عدد المصابيح			
[3,4[	3,5	15	15	0.0375	3,75%
[4,5[	4,5	46	61	0.115	11,5%
[5,6[	5,5	54	115	0.135	13,5%
[6,7[	6,5	78	193	0.195	19,5%
[7,8[	7,5	70	263	0.175	17,5%
[8,9[	8,5	64	327	0.16	16%
[9,10[	9,5	45	372	0.1125	11,25%
[10,11[	10,5	20	392	0.05	5%
[11,12[	11,5	8	400	0.02	2%

IV) التعبير البياني والرسوم المبيانية:

يخص التعبير البياني، مبادئ الكارطوغرافيا البيانيـة وتقنياتهـا الأسـاسـية. والتركيـز هنـا يكون على البيانـات الإحصـائية، لأن الخرائطيـة تُـدرس فـي مـواد الخـرائط، انطلاقـا مـن الإحداثيات في المستوى والفضاء حيث يتم إنشاؤها في بُعدين أو ثلاثة أبعاد.

أمثلة كارطوغرافية: تقنيات سُلم التصاميم والخرائط،...

تمرين: أنشئ نقطا بإحداثيات في المستوى والفضاء.

و في مجال الإحصاء، تُطبق تلك المبادئ على الرسوم المبيانيـة، المتعـددة والمتنوعـة حسـب الحالات الإحصائية، إلا أنها صارت متيسـرة بفضل تطور الإعلاميات الحديثة.

كما أنها يدويا، تُنشأ في معلم متعامد، ليس بالضرورة ممنظما ولا مرتبا دائما من أصله. فيُستعمل سُلم لتبسيط الرسم، بل قد تُرسم الترددات المئوية فقـط إذا كانـت الأعـداد كبيرة، مع أخْذ نِسـب صحيحة مثل التكرار ومناسبة بإفراط أو تفريط لكـي يبقـى التكـرار الإجمالي 100.

و من بين **الرسوم المبيانية**، ما يلي:

+ **مخطط بالقضبان**: في معلم متعامد أو مربع تُمثل فيه النسب.

+ **مخطط عصوي**: تُسمى أعمدة أيضا مثل القضبان العمودية.

+ **مدراج**: درجات.

+ **المضـلع الإحصـائي**: يحـدد المسـاحة نفسـها المحـددة بالمـدراج المتسـاوي المجالات، وإلا وجب تعديل جدول الفئات.
+ **مخطط دائري Pi * 3,6 ، أو نصف دائري Pi * 1,8:** وهذا الأخير، يصلح للمقارنة بين متغيرين، من خلال نصفي دائرة متماثلين مع قُطر مختلف.
+ **الهرم (Pyramide):**
يصلح للمقارنة بين متغيرين، من خلال قضبان متماثلة أفقيا، وفئات نفسها مع تكرارات ترتبت تنازلية كلها أو أكثريتها. فيقع رسم شبيه بالهرم، قمته في الأعلى.
مثال للهرم: في مجموعة للطلبة، كانت النتائج كالآتي:

مجال المعدل	الذكور	الإناث
أقل من 9	15	17
[9,11[	20	26
[11,13[	15	24
[13,15[	10	13

V) الوسيطات:
خاصة بالمتسلسلات الإحصائية الكمية، وتلحق بها النوعية التي يمكن تحويلها إلى كمية، كالنتائج القابلة للترتيب الرقمي والتواريخ المتتابعة.
أ) مقاييس (وسيطات) النزعة المركزية:
1) المنوال (Mo):
المواصفة التي لها أكبر تكرار والتكرار الأعلى. فيمكن أن يوجد أكثر من منوال، لكن لا يعتد بأكثر من منوالين، وكذلك الشأن في الصنف المنوالي.
+ مثال 2: أكبر تكرار هو 21، المنوال هو 0.
+ مثال 3: أكبر تكرار هو 78، الصنف المنوالي هو [6,7[.
ملاحظة: في حالة الصنف المنوالي الوحيد [a , b[يعتبر المنوال Mo تقريبا.

$$Mo = a + \underbrace{(b - a)}_{\text{طول الفئة}} \times \frac{P}{P + Q}$$

وهناك طريقة اخرى تسمى الرافعة باعتبار P التكرار اللاحق و Q السابق.
كما يُستعمل رسم الصنف المنوالي مع سابقه ولاحقه بتقاطع الخطين، الجامع بين الحدود العليا للمنوالي وسابقه والجامع بين الحدود الدنيا للمنوالي ولاحقه
P: الفرق بين تكرار الصنف المنوالي وتكرار الصنف الذي قبله.
Q: الفرق بين تكرار الصنف المنوالي وتكرار الصنف الذي بعده.
تطبيق: $Mo = 6 + (7-6) \times \frac{24}{24+8} = 6,75$

2) المعدلات:
متناسبة فيما بينها، إلا أن كلا منها يناسب نوعا أكثر من غيره. فكان من أهمها:
2 – 1) الوسط الحسابي:

الذي قد يرمز له بـ $\bar{x}$، من مجموعة اعتبارية، وبـ m من ملاحظة جزئية بجوار المجموعة، وبـ μ حقيقة أو على الأقل نظريا.

و هو $\sum xi\ fi = \frac{\sum_{i=1}^{n} xi * ni}{N}$.

و في حالة الأصناف، يكون مركز الصنف مكان المواصفة. لكن عند عدم وجود مركز، كما في أكبر أو أصغر من قيمة معينة، فإنه يؤخذ الطرف المُنته.

مثال 2: الوسط الحسابي هو:

xi ni	ni	xi
0	21	0
18	18	1
14	7	2
9	3	3
4	1	4

$\sum = 45 / 50 = 0, 9$

ملاحظات:

1) هذا المعدل الأكثر تداولا، غير واقعي دائما، كما في المثال بعدد غير صحيح عن الحوادث، إذ إنه تقريبي فقط.

2) الوسط الحسابي لسلسلة غير متزنة من p قيمة غير متكررة، هو $\sum xi / p$.

3) معدل المعدلات:

إذا وُزعت الساكنة الإحصائية إلى أجزاء منفصلة، ذات تكرار إجمالي جزئي Ni ومعدل mi، فإن معدل المعدلات الجزئية هو الوسط الحسابي، أي $\frac{\sum mi\ Ni}{N}$

مثال: في قسم، معدل نقط الطلبة العشرين هو 12 ومعدل الطالبات الثلاثين هو 13. إذن الوسط الحسابي الإجمالي هو:

$$\frac{20 \times 12 + 30 \times 13}{50} = 12,6$$

تمرين: في منطقة، معدل درجة الحرارة في الشهور الثلاثة الباردة هو 8، وفي الخمسة الحارة هو 35، وفي المعتدلة 20. فما هو المعدل السنوي للحرارة ؟.

2 – 2): المعدل الهندسي: المرموز له بـ G هو: $\sqrt[N]{\Pi xi^{ni}}$

و في حالة الأصناف، يكون مركز الصنف مكان قيمة الميزة.

ملاحظات:

1) المعدل الهندسي لسلسلة غير متزنة، من p قيمة غير متكررة، هو الجذر من الرتبة p لجداء القيم.

2) يُحسب باللوغاريتم، إلا أن الحساب المباشر متيسر آليا ومع جدول.

3) هذا المعدل أفضل في حالة النسب، مثل التغيير بدْءا من قيمة أصلية V_0، والنمو السكاني مع نِسب مختلفة كما في الغالب. فاختير معدل للنسب هندسي p، يحقق:

$V_n = V_0 (1 + p)^n$ أو $V_n = (1+p_1)\ldots(1+p_n)\ V_0$

و $(1 + p)^n = (1 + p_1) \ldots (1 + p_n)$، مع كتابة النسب p_i عشريا،أي: $p = \sqrt[n]{V_n/V_0} - 1$

تمرين:

نمو مائة ألف من السكان خلال أربع سنوات، كان على التوالي:

20%، 20%-، 30%، 25%.

حدد القيمة النهائية ومعدل نسبة النمو السنوي.

2 – 3) المعدل التوافقي: المرموز له بـ H يحقق: $\frac{1}{H} = \frac{\sum_{i=1}^{n} \frac{ni}{xi}}{N}$.

ملاحظات:

1) في المعدل التوافقي غير المتزن، يكون التكرار 1.
2) الحساب المباشر متيسر آليا ومع جدول.
2) هذا المعدل أفضل في حالات، مثل معدلات الأسعار والسرعة.
3) $H \le G \le m$.

مثال:

2 –4) المعدل التربيعي: المرموز له بـ Q يحقق: $Q^2 = \frac{\sum_{i=1}^{n} xi^2 * ni}{N}$.

فكان أقل استعمالا، لأنه يُحسب بالجذر المربع، رغم أنه يقارب المعدل الحسابي.
مثال سابق للمقارنة:

3) الوسيط:

المرموز له بـ M يحقق مجموع تكرارات القيم الأصغر منه ومجموع تكرارات القيم الأكبر منه، لا يفوقان نصف التكرار الإجمالي. فيكون نصف الوحدات على الأقل، بقيم أصغر من أو تساوي M، ونصف الوحدات على الأقل أيضا، بقيم أكبر من أو تساوي M.

3 – 1) طريقة لتحديده:

*** في حالة الكمية المتقطعة:**

هو أصغر قيم المتغير الإحصائي، التي تكرارها المتراكم، أكبر من أو يساوي نصف التكرار الإجمالي. وفي حالة التساوي، تكون قيمتان وسطيتان متتاليتان.

مثال 2:

نصف التكرار الإجمالي هو 25، والتكرار المتراكم الأكبر منه أو يساويه هو 39.
إذن الوسيط هو الميزة المقابلة، أي 1.

تمرين: حدد الوسيط

9	7	6	3	1	0	x_i
1	1	3	1	2	2	n_I

*** في حالة الكمية المتصلة (الصنف الوسطي يضم M):**

$N_{k-1} \le N/2 < N_k$ و $a_{k-1} \le M < a_k$ فإن: $\frac{a_k - a_{k-1}}{N_k - N_{k-1}} = \frac{M - a_{k-1}}{\frac{N}{2} - N_{k-1}}$

مثال 3:

نصف التكرار الإجمالي هو 200، والتكرار المتراكم الأكبر منه أو يساويه هو 263، الموافق للمجال]7,8[، الذي يضم الوسيط M:

إذن $7 \le M < 8$

$193 \le 200 < 263$

فنحسب بالطريقة الثلاثية: $\frac{8-7}{263-193} = \frac{M-7}{200-193}$ أي $M - 7 = \frac{7}{70}$ ثم $M = 7 + \frac{1}{7} = 7.14$

ملاحظات:

1) M محصور بين m و M_O، الوسط الحسابي والمنوال.
2) في حالة متسلسلة غير متزنة، يُلاحظ الوسيط مباشرة وفي الوسط، حسب زوجية عدد القيم.

3) يتحدد الوسيط مبيانيا، لأنه أفصول نقطة تقاطع المضلعين الإحصائيين للتكرارات، التصاعدية والتنازلية.
4) في كمية متصلة، يُؤخذ $^{N}0 = 0$ عندما يقع: $0 \le {}^{N}/2 < {}^{N}1$ و $^{a}0 \le M < {}^{a}1$. كما يُتفادى الوقوع في مقام منعدم، باختيار الفروق غير المنعدمة.
5) تُحسب M أيضا بالترددات مع 0,5 و النسب مع 50%، بدلا من التكرارات المتراكمة مع $^{N}/2$. فكان الوسيط مختلف الطرق، إلا أنها قليلة الاستعمال، لعدم خضوعها لعمليات حسابية دقيقة عوض طريقة التأطير.
6) إذا كان <u>الوسيط</u> أحد مقاييس النزعة المركزية، فإن هناك مقياسا مشابها يُرمز له بـ ^{M}L (médiale)، إلا أنه لتمركز وكتلة قيم المتغير الإحصائي مع تكراراتها. فيكون أكبر من أو يساوي M، إلا أنه يُحسب مثلها سواء في المتقطع أو المتصل، مع اعتبار المجموع المتراكم للجداءات xi*ni ونصف المجموع الإجمالي $\frac{\sum_{i=1}^{n} xi * ni}{2}$.

<u>أمثلة</u> نفسها للمقارنة بين M و^{M}L:

<u>3 – 2) تمديد الوسيط:</u>

إذا كان الوسيط يقسم المتسلسلة الإحصائية إلى جزئين متساويين، فإن:
+ الربيعات Qi (تكسيرات من الرتبة 4: quartiles): تقسمها إلى أربعة أجزاء متساوية، لتوافق $\frac{iN}{4}$.
+ العشيرات Di (تكسيرات من الرتبة 10: déciles): تقسمها إلى عشرة أجزاء متساوية، لتوافق $\frac{iN}{10}$.
+ المؤينات Ci (تكسيرات من الرتبة 100: centiles): تقسمها إلى مائة جزء بالتساوي، لتوافق $\frac{iN}{100}$.

<u>ملاحظات:</u>
1) هذه التجزيئات، تُحدد حسابيا مثل الوسيط، مع استعمال ما يوافقها، وهناك أيضا غيرها.
2) $M = Q_2 = D_5 = C_{50}$.

<u>أمثلة</u> في المتقطع والمتصل:

<u>تمرين:</u> بين أن للمتسلسلتين التاليتين، مقاييس التموضع نفسها، على الرغم من اختلافهما.

xi	8	9	10	11	12	13	14
ni	1	2	3	5	3	2	1

yi	2	5	7	8	10	11	12	14	15	17	20
ni	1	2	1	2	1	3	1	2	1	2	1

<u>استنتاج:</u>

مقاييس التموضع مهمة بالمعلومات المستنبطة منها، إلا أنها غير كافية لتمييز متسلسلة إحصائية. فكان ضروريا تحديد وسيطات التشتت، التي تضبط مدى تشتت قيم المتسلسلة حول مركزها، لمعرفة مستوى تجانس وانسجام وتمركز القيم.

<u>ب) مقاييس (وسيطات) التشتت:</u>

1) مدى المتسلسلة:

هو الفرق بين أعلى وأصغر قيم المتغير الاحصائي. فيتحدد مجال التغيير الحقيقي، أو الاختياري كما في أطراف الأصناف المختارة. لكنه مقياس غير مفيد دائما، كما في حالة وجود قيم متطرفة.

مثال التمرين: الفارق المطلق في xi هو 6 = 8 - 14، وفي yi هو 18 =2 - 20.

ملاحظة: في حالة لزوم تحديد فئات، فإن طول الفئة هو: $\frac{\text{الفارق المطلق} + 1}{\text{عدد الفئات}}$

وهذا العدد المطلوب، قد يكون مختارا فقط مناسبا، أو ينبني على معادلات محددة، مثل: $1 + 3,332 \log N$ و $2,5\sqrt[4]{N}$.

مثال xi: نفترض الحصول على ثلاثة فئات، ليكون طول الفئة هو:

$\frac{6+1}{3} = \frac{7}{3} = 2,33$.

فنعتبر طول الفئة هو 3، ليصبح الجدول:

الفئة	التكرار
[8-11[	6
[11-14[	10
[14-17[	1

2) الانحرافات التكسيرية: هي أكثر دقة، في ضبط التشتت حول المركز لقيم المتغير الإحصائي.

+ **الفارق المطلق** /الانحراف الربيعي للنسبة 50% écart interquartile هو Q_3-Q_1.
+ **الفارق المطلق** /الانحراف العشيري للنسبة 80% écart interdécile هو D_9-D_1.
+ **الفارق المطلق** /الانحراف المؤيني للنسبة 98% écart intercentile هو C_{99}-C_1.

مثال:

3) الفارق المطلق المتوسط (تشتت حول الوسط الحسابي):

$$e = \frac{\sum ni\,|xi - \bar{x}|}{N}$$

أمثلة:

4) التباين: المرموز له بـ V أو σ^2 أو σ'^2 أو S^2 ، هو: $V = \frac{\sum ni\,(xi - \bar{x})^2}{N}$

طريقة ثانية: $V = \frac{\sum ni\,xi^2}{N} - \bar{x}^2$.

أمثلة:

ملاحظة: يُستعمل تباين مُجَمّع، لمقارنة مجموعتين X ' و'' X من نفس المجتمع الإحصائي، محققة: $V^2 = \frac{(N'-1)\,V'^2 + (N''-1)\,V''^2}{N' + N'' - 2}$.

فيكون (2 -) من مجموعتين، و (3 –) من ثلاثة مع إضافة المجموعة الثالثة،...

5) الانحراف الطرازي: المرموز له بـ σ أو σ' أو S هو $\sqrt{V}$.

ملاحظات:

1 – في حالة الفئات، يكون مركز الصنف مكان قيمة المتغير الإحصائي، إلا إذا وجدت فئات متطرفة غير محدودة، حيث يؤخذ الطرف المنته.

2 – الفارق المطلق المتوسط، لا يُستعمل كثيرا، بسبب تَعقد حساب القيم المطلقة. فاستُبدلت بالمربعات في التباين، خاصة أنها أسهل في طريقتها الثانية. وبما أن المربع يؤدي إلى اختلاف، كما في وحدة القياس، فقد استُعمل الجذر التربيعي في الانحراف الطرازي للمحافظة على الوضع نفسه. فكان هذا الانحراف الأخير، هو أحسن تلك الوسيطات والمقاييس.

ج) المعاملات الإحصائية:

1) معامل الاختلاف:

المرموز له بـ CV هو 100 % * $\frac{\sigma}{m}$. مدى التشتت حول الوسط الحسابي.

إذا كان CV < 2%، فإن التشتت ضعيف جدا لعوامل عكس حالات أخرى.
و إذا كان CV > 5%، فإن هناك خللا ووضعا غير عادي، فيلزم تعديل موضع الدراسة الإحصائية لتكون أقرب إلى الصواب.
و في حالة ما إذا كان CV > 20% ، فإن الدراسة تكون غير طبيعية التشتت.
فيحدد هدا المعدل، مدى صلاحية المجموعة والدراسة الإحصائية. كما يقيس التشتت النسبي، ويصلح للمقارنة خاصة أنه مقياس دون وحدة القياس.
و في حالة وجود مجال غير محدود، يُستعمل:

معامل الاختلاف الربيعي: $CV = \frac{Q3-Q1}{Q3+Q1} * 100\%$

و قد تستعمل صيغ أخرى، مثل: $CV = \frac{Q3-Q1}{M} * 100\%$

2) معامل الالتواء:

الذي يُرمز له بـ CA هو $CA = \frac{m - Mo}{\sigma}$ (مقياس بيرسون)

ليقيس الفرق بين الوسط الحسابي والمنوال، بالنسبة إلى الانحراف σ .
ففي حالة التوزيع الطبيعي المتماثل، تتساوى مقاييس التموضع الأساسية:
m=M=Mo (المنوال، الوسط الحسابي، الوسيط).
لكن في الغالب غير المتماثل، يكون التوزيع من قِمّته مائلا إما يمينا أو يسارا.
فيكون ملتويا يسارا، إذا كان موجب الالتواء حيث الوسط الحسابي أكبر من المنوال قطعا ($m_o < M < m$).
ويكون ملتويا يمينا، إذا كان سالب الالتواء حيث الوسط الحسابي أصغر من المنوال قطعا ($m < M < mo$).
وبقدر الابتعاد عن 0، يكون الالتواء أكثر حدة، مع العلم أن $-3 \leq CA \leq 3$ (القرب او البعد عن الاعتدال).
و هناك صيغ أخرى مغايرة أحيانا، كما في حالة توزيع قريب من المتماثل، حيث يكون الفرق بين الوسط الحسابي والمنوال صغيرا جدا، فضُرب في العدد 3، ليصير معامل الالتواء: $CA = 3 \frac{m - Mo}{\sigma}$

ثم في حالة مجالات متطرفة غير محدودة، هناك صيغة مبنية على كون:
M-Q1 وQ3-M يحددان نفس النسبة 50% من النصف، ليصير معامل الالتواء:
$CA = 1 - \frac{2M}{Q3-Q1} = \frac{(Q3-M) - (M-Q1)}{Q3-Q1}$ (مقياس بول)

3) معامل التمركز:

المرموز له بـ CC ، يحدد درجة اللاتكافؤ في توزيع قيمة إحصائية إجمالية بين وحدات مجتمع. ويخص بالذات، التوزيعات التكرارية المبوبة على شكل فئات، لأنها تقريبية فقط. فيوجد معطى آخر دقيق، هو القيمة الإجمالية للفئة، أي مجموع القيم.

وإذا لم يُعط، يعوَّض بجداء مركز الفئة في حصيصها التكراري. فيكون منحنى أو مضلع التمركز، الذي يربط النسبة التراكمية مع نِسب قيم الفئات الإجمالية التراكمية. ويتم مربع بين (0 , 0) و(100 , 100)، يحقق قطره التوزيع المثالي العادل تماما، حيث التساوي مع القيمة الإجمالية الحقيقية للفئة في الواقع.

وكلما ابتعد منحنى التمركز عن الوَتر تحته حتما، كلما كان التمركز شديدا. فيقاس بحساب المساحة الواقعة بين المنحنى والوتر، أي الفرق بين مساحة نصف المربع (5000 = 2 / 100 * 100) ومجموع المساحات تحت المنحنى (مثلث وأشباه منحرف).

ومعامل التمركز، هو نسبة الفرق على 5000. فيكون التمركز حادا، إذا كان المعامل أكبر من 0 , 5 أي 50%.

مثال:

فئات القيم	التكرار	القيمة الإجمالية	النسبة %	نسبة القيمة الإجمالية	النسبة التراكمية	نسبة القيم الإجمالية تراكمية	المساحات تحت منحنى التمركز
[2,4[	24	60	12	4,5	12	4,5	27
[4,6[	50	220	25	16,4	37	20,9	317,5
[6,8[	80	560	40	41,8	77	62,7	1672
[8,10[	30	280	15	20,9	92	83,6	1097,25
[10,12[	16	220	8	16,4	100	100	734,4
$\sum$	200	1340	100	100			3848,15

كيفية حساب المساحات:

المثلث: $\frac{1}{2}27 = (12 * 4,5)$.

شبه المنحرف الأول: $\frac{1}{2}25 * (20,9 + 4,5)$ ، حيث 12 – 37 = 25.

شبه المنحرف الثاني: $\frac{1}{2}40 * (62,7 + 20,9)$ ، حيث 37 – 77 = 40.

و هكذا...

$CC = \frac{5000 - 3848,15}{5000} \approx 0,23 = 23\ \%$: تمركز معتدل.

تمرين: مثال سابق بفئات، لتُحدد القيم الإجمالية بالجداء.

ملاحظة: $0 \leq cc \leq 1$

Cc=0 ←انعدام التمركز أي توزيع متكافئ

Cc=1 ← s=5000أي تمركز تام

Ccقريب من 0 ←تمركز ضعيف

Cc قريب من 1 ←تمركز قوي

الارتباط (corrélation) والانحدار (regression):

1) الارتباط r:

المقصود هو قياس درجة العلاقة والترابط والتأثير والتأثر بين متغيرين، فيسمى ارتباطا بسيطا، أو بين أكثر من اثنين، فيسمى ارتباطا متعددا. وكما يكون كميا عدديا، قد يكون نوعيا، خاصة مع الرُّتب.

خصائص:

1) $-1 \le r \le 1$ فيحسب بقاعدة موجبة ثم تحدد اشارته بملاحظة مدى مسايرة قيم متغير لقيم متغير اخر.

2) الانحراف المعياري لمعامل الارتباط، الذي يصلح للمقارنة والتمييز بين الارتباطات عند الاقتضاء، هو: $s\ r = \sqrt{\frac{1-r^2}{n-2}}$، حيث n عدد القيم الإجمالي.

3) كلما اقترب r من 1 أو -1، كلما كان الارتباط قويا، إلا أن الطّردي يكون موجبا، كتزايد انجراف التربة مع تزايد كمية الأمطار، في حين أن العكسي يكون سالبا، كانخفاض الحرارة مع ارتفاع العلو. وكلما اقتربا من 0، كلما كان الارتباط ضعيفا، إلى أن يصل إلى الاستقلال التام.

أ) الارتباط النوعي: بين متغيرين نوعيين مرتبين بـ X i وYi، محدد بالمعامل:

$r = 1 - \frac{6\sum_{i=1}^{n} di^2}{n(n^2-1)}$، حيث d i الفرق بين رتبتي Xi وYi.

مثال:

سرعة الرياح Xi	تكون كثبان الرمل Yi	رتبة Xi	رتبة Yi	di	di ²
كبيرة	ضخمة	1	1	0	0
متوسطة	متوسطة	2	2	0	0
ضعيفة	ضئيلة	3	3	0	0
ضعيفة جدا	ضئيلة جدا	4	3	1	1

$r = 1 - \frac{6*1}{4(4^2-1)} = 0,9$، أي ارتباط قوي نسبته 90 %، و$\sqrt{\frac{1-0,9^2}{4-2}} \approx 0,38$ = sr.

ب) الارتباط الكمي: بين متغيرين كميين متقطعين x i وyi، وإلا استُعمل مركز الصنف. فكان محددا بالمعامل: $r = \frac{\sum_{i=1}^{n} xi * yi}{\sqrt{\sum_{i=1}^{n} xi^2 \sum_{i=1}^{n} yi^2}}$.

مثال: العلاقة بين عدد المراكز الصحية وميزانية عمالات مقاطعات الدار البيضاء، حسب إحدى الإحصائيات:

العمالات	الميزانية بمليون درهم xi	عدد المراكز الصحية yi	xi * yi	xi²	yi²
عين الشق	9	17	153	81	289

عين السبع	13.7	14	191.8	187.69	196
الفداء	49	17	833	2401	289
بن امسيك	17.7	8	141.6	313.29	64
آنفا	15	15	225	225	225
المشور	42	2	84	1764	4
م.رشييد	10	9	180	400	81
البرنوصي	21	7	147	441	49
$\sum$	187.4	89	1955.4	5812.98	1197

$r = \frac{1955.4}{\sqrt{5812.98*1197}} \approx 0.74$ ، أي ارتباط لا بأس به، و$s\,r = \sqrt{\frac{1-0,74^2}{8-2}} \approx 0,28$.

<u>2) الانحدار المستقيم البسيط:</u>

يقيس العلاقة النسبية y/x بين متسلسلتين غير متزنتين xi وyi، لتحديد درجة تحكم x في y. ويرتكز على مستقيم بمعادلة y = a x + b، حيث الميل أو معامل الانحدار هو $a = \frac{\sum_{i=1}^{n}(xi*yi) - n\bar{x}\bar{y}}{\sum_{i=1}^{n}(xi^2) - n\bar{x}^2}$ ، وتقدير نسبة الانحدار في قيمة $b = \bar{y} - a\bar{x}$.

<u>ملاحظات:</u>

1) $\frac{\text{cov}(xi, yi)}{V(xi)} = a = \frac{\frac{\sum_{i=1}^{n} xi*yi}{n} - \bar{x}\bar{y}}{\frac{\sum_{i=1}^{n} xi^2}{n} - \bar{x}^2}$.

2) بالمثل يُعَرف انحدار x/y .

<u>مثال:</u> السابق.